Stochastic Uncertainty Analysis for Solute Transport in Randomly Heterogeneous Media Using a Karhunen–Loeve Based Moment Equation Approach

Gaisheng Liu[1], Zhiming Lu[2], and Dongxiao Zhang[1,3]
[1]Mewbourne School of Petroleum and Geological Engineering, University of Oklahoma, Norman, OK 73019
[2]Hydrology, Geochemistry, and Geology Group, Los Alamos National Laboratory, Los Alamos, NM 87545
[3]Department of Energy and Resources Engineering, College of Engineering, Peking University, Beijing 100871, P. R. China

Abstract

The Karhunen-Loeve based moment equation (KLME) approach proposed by *Zhang and Lu* [2004] has been extended to solving solute transport problems in randomly heterogeneous media. The KLME approach combines the Karhunen-Loeve decomposition of the underlying random conductivity field and the perturbative and polynomial expansions of other dependent variables including the hydraulic heads, flow velocities, dispersion terms and solute concentrations. The equations obtained in this approach are sequential and their structure is formulated in the same form as the original governing equations such that any existing simulator, such as MT3DMS, can be applied as the solver. Through a series of 2-D examples, the validity of the KLME approach is evaluated against the classical Monte Carlo (MC) simulations. Under all different flow and transport conditions considered in this work, the KLME approach provides an accurate representation of the mean concentrations. For concentration variances, the accuracy of KLME approach is reasonably good at a small conductivity variance. However, as the conductivity variance increases from 0.25 to 0.5, the mismatch between KLME and MC results becomes large. Our results also indicate that when the local dispersivities are relatively large, neglecting the effects of the cross terms between velocity fluctuations and local dispersivities, as commonly

done in previous studies, produces large errors on the concentration variances and a rigorous treatment of the dispersion terms is more appropriate.

1. Introduction

Subsurface fluid flow and transport processes often take place in a complex geologic setting where permeability, the key medium property in controlling flow and transport behaviors, exhibits a high degree of spatial variability and cannot be accurately characterized in all the details. As a result, our model predictions of flow velocities and solute concentrations are subject to a great deal of uncertainty. To address uncertainty in subsurface flow and transport modeling, stochastic approaches have been developed [*Dagan*, 1989; *Gelhar*, 1993; *Cushman*, 1997; *Zhang*, 2002; *Rubin*, 2003]. Early stochastic transport research has emphasized on reproducing the ensemble averaged plume behaviors using effective macrodispersion coefficients [*Gelhar et al.*, 1979; *Dagan*, 1984; *Neuman and Zhang*, 1990]. Although the macrodispersion approach provides a reasonable representation of the gross spatial spreading of solutes due to random heterogeneity, it does not provide an estimation of the uncertainty associated with the mean predictions (e.g., fluctuations around mean concentrations or concentration variances). Since the early 1990s, concentration variance has become a subject of great interest in various studies [*Dagan et al*., 1992; *Neuman*, 1993; *Kapoor and Gelhar*, 1994a, b; *Zhang and Neuman*, 1996; *Dagan and Fiori*, 1997; *Morales-Casique et al.*, 2006a, b]. It has been demonstrated that while local-scale dispersion has a relatively insignificant impact on the mean concentration predictions as compared to the field-scale heterogeneities, it can exert a strong influence on the concentration variance results.

Stochastic approaches can be generally sorted into two different frameworks, namely, Monte Carlo (MC) simulations and the moment equation (ME) approaches. Since the MC method is a brute force approach, it is computationally demanding for large-scale problems and it typically serves as a benchmark model for evaluating the accuracy of other approaches [*Graham and McLaughlin*, 1989; *Hassan et al*., 1998]. In the ME approaches the statistical moment equations are directly derived for model predictions using the perturbation technique [*Neuman*, 1993; *Kapoor and Gelhar*, 1994a, b; *Zhang and Neuman*, 1995, 1996; *Hu et al*., 1999; *Morales-Casique et al.*, 2006a, b]. To mitigate the closure problems caused by neglecting the triplet term in the macrodispersive flux in the Eulerian approach, *Hu et al*. [1999] expanded the concentration as an infinite series instead of the mean and perturbation decomposition. The mean concentrations and spatial moments were solved in the Fourier-Laplace space up to different expansion orders in flow and transport, respectively. *Morales-Casique et al.* [2006a, b] present the first and second MEs for advective-dispersive transport and propose a higher-order iterative closure scheme for the special case of steady state flow with respect to the first order in log conductivity variance. Compared to the MC simulations, the perturbative moment solutions of stochastic flow and transport problems are formally limited to mild medium variability although data conditioning can certainly increase the effective range to some extent. More recently, *Neuman* [2006] suggests combining fractal and variational multiscale decomposition in order to extend the applicability of perturbative ME approaches in composite media where the heterogeneity can be arbitrarily large. In general, analytical solutions of the moment equations can be obtained with the aid of Green's function for some limiting cases under simplified conditions. Numerical ME approaches are conceivable, but the computational effort increases rapidly with the size of the problem, thus limiting its applicability to large-scale problems.

In recent years a new class of stochastic approaches have been developed that rely on the Karhunen-Loeve (KL) decomposition of the underlying random fields [*Ghanem and Spanos*, 1991; *Ghanem and Dham*, 1998; *Zhang and Lu*, 2004]. Unlike the conventional ME ones, the KL-based approaches do not require solving the covariance and cross-covariance matrices directly and thus become more efficient computationally. *Ghanem and Dham* [1998] combined the KL decomposition of random intrinsic permeability field and orthogonal polynomial chaos expansions of other stochastic dependent variables and applied the KL/polynomial chaos method to a 2-D multiphase flow problem. *Zhang and Lu* [2004] proposed to integrate the KL decomposition with perturbative and polynomial expansions of other stochastic dependent variables, and the resulting methodology was referred to as KLME. In contrast to the KL/polynomial chaos method where equations of different orders are interactively coupled, the equations in the KLME approach are recursive and can be solved sequentially from low to high orders. The KLME approach has been applied to different types of problems including conditional simulations, saturated-unsaturated, steady-state two-phase and unconfined flow [*Lu and Zhang*, 2004; *Yang et al*., 2004; *Chen et al*., 2005; *Liu et al*., 2006].

Motivated by the recent success in its flow applications, we extend the KLME approach to stochastic solute transport problems. Both the mean concentrations and concentration variances are estimated in this work. Due to its importance, local dispersion is included and implemented in a rigorous manner. Previous studies [*Kapoor and Gelhar*, 1994a, b; *Zhang and Neuman*, 1996; *Hu et al.*, 1999; *Morales-Casique et al.*, 2006a, b] have treated local dispersion as either given constants or a linear function of mean flow velocities. In this work we shall demonstrate that this simple approximation has somehow underestimated the heterogeneity effects on dispersion and can produce inappropriate results under certain circumstances.

The remainder of paper is organized as follows. We begin by describing the mathematical model that we solve for solute transport under advection, dispersion and external sinks/sources. Next, we present the theoretical derivations of stochastic transport formulations using the KLME approach and describe their solution procedure. Finally, we evaluate the validity of KLME approach in a series of 2-D examples by comparing to the MC simulation results.

2. Methodology

2.1. Mathematic Model

The transport of a conservative solute in 3-D groundwater flow is given by the advection-dispersion equation with possible external sinks/sources,

$$\frac{\partial C(\boldsymbol{x},t)}{\partial t} = \nabla \cdot \left(D(\boldsymbol{x},t)\nabla C(\boldsymbol{x},t)\right) - \nabla \cdot \left(v(\boldsymbol{x},t)C(\boldsymbol{x},t)\right) + q_s C_s , \tag{1a}$$

subject to following initial and boundary conditions,

$$C(\boldsymbol{x},0) = C_0(\boldsymbol{x}), \qquad \boldsymbol{x} \in \Omega , \tag{1b}$$

$$C(\boldsymbol{x},t) = C_D(\boldsymbol{x},t), \qquad \boldsymbol{x} \in \Gamma_D , \tag{1c}$$

$$D(\boldsymbol{x},t)\nabla C(\boldsymbol{x},t) \cdot \boldsymbol{n}(\boldsymbol{x}) = -F(\boldsymbol{x},t), \qquad \boldsymbol{x} \in \Gamma_N , \tag{1d}$$

where C is the solute concentration; D is the hydrodynamic dispersion tensor; v is the pore water velocity vector $(v_x(\boldsymbol{x},t), v_y(\boldsymbol{x},t), v_z(\boldsymbol{x},t))^T$ (where superscript T indicates transpose); q_s and C_s are the flow rate and solute concentration in the sinks/sources; $\boldsymbol{x}$ is the vector of spatial Cartesian coordinate $(x, y, z)^T$; t is time; C_0 is the initial concentration in the transport domain Ω; $C_D(\boldsymbol{x},t)$ is the specified concentration on the Dirichlet boundary segments Γ_D; $F(\boldsymbol{x}, t)$ is the dispersive flux across Neumann boundary segments Γ_N; and $\boldsymbol{n}(\boldsymbol{x})$ is an outward unit vector normal to the boundary $\Gamma_D \cup \Gamma_N$. For simplicity, the spatial and temporal indices $\boldsymbol{x}$ and t are omitted in the remainder of paper.

The hydrodynamic dispersion tensor, D, in a locally isotropic medium, with an accommodation made for different orthogonal transverse dispersivity values, can be expressed as [*Burnett and Frind*, 1987],

$$D_{xx} = \left(\alpha_L v_x^{\,2} + \alpha_{TH} v_y^{\,2} + \alpha_{TV} v_z^{\,2}\right)/|v| + D^*,$$

$$D_{yy} = \left(\alpha_L v_y^{\,2} + \alpha_{TH} v_x^{\,2} + \alpha_{TV} v_z^{\,2}\right)/|v| + D^*,$$

$$D_{zz} = \left(\alpha_L v_z^{\,2} + \alpha_{TH} v_x^{\,2} + \alpha_{TV} v_y^{\,2}\right)/|v| + D^*,$$

$$D_{xy} = D_{yx} = (\alpha_L - \alpha_{TH}) v_x v_y / |v|, \qquad (2)$$

$$D_{xz} = D_{zx} = (\alpha_L - \alpha_{TV}) v_x v_z / |v|,$$

$$D_{yz} = D_{zy} = (\alpha_L - \alpha_{TV}) v_y v_z / |v|,$$

where v_x, v_y, and v_z are the components of the pore water velocity v and $|v|$ is its magnitude; α_L is the longitudinal dispersivity; α_{TH} and α_{TV} are the transverse dispersivities in the horizontal and vertical directions, respectively, and D^* is the molecular diffusion coefficient in porous media.

Due to the uncertainty associated with the hydraulic conductivity (or permeability), the hydraulic head and pore water velocity become stochastic, and so does the concentration of solute that is moved by flow. The external sink/source term $q_s C_s$ is assumed to be deterministic. The objective of this work is to solve for the mean solute concentrations and the errors associated with mean predictions through concentration variances. Stochastic solutions of flow problems using the KLME approach can be found in our previous work [e.g., *Zhang and Lu*, 2004; *Lu and Zhang*, 2004; *Liu et al.*, 2006] and are thus not repeated here.

2.2. Karhunen-Loeve Expansion-Based Moment Equations (KLME)

To solve (1) using the KLME approach, we first expand the stochastic variables C, D, v as an infinite series,

$$C = \sum_{m=0}^{\infty} C^{(m)}, \quad D = \sum_{m=0}^{\infty} D^{(m)}, \quad v = \sum_{m=0}^{\infty} v^{(m)}, \tag{3}$$

where $C^{(m)}$, $D^{(m)}$ and $v^{(m)}$ are the m^{th}-order expansions with respect to the standard deviation of log hydraulic conductivity, σ_Y. The detailed expressions of $v^{(m)}$ can be found in *Lu and Zhang* [2004]. The derivations of $D^{(m)}$ are provided in Appendix A for m up to the third order. It is noteworthy that in previous studies [*Kapoor and Gelhar*, 1994a, b; *Zhang and Neuman*, 1996; *Hu et al*., 1999; *Morales-Casique et al.*, 2006a, b], the dispersion tensor D has been treated as given constants or linear functions of mean flow velocities, and therefore the effects of velocity variations on D are disregarded arbitrarily. As will be demonstrated in the following sections, this approximation, equivalent to using $D^{(0)}$ to approximate D in (3), may become problematic in certain cases where higher-order $D^{(m)}$ arising from the velocity fluctuations are significant.

Substituting (3) into (1a) and rearranging the summations, one obtains,

$$\sum_{m=0}^{\infty} \frac{\partial C^{(m)}}{\partial t} = \sum_{m=0}^{\infty} \sum_{k=0}^{m} \nabla \cdot \left(D^{(k)} \nabla C^{(m-k)} \right) - \sum_{m=0}^{\infty} \sum_{k=0}^{m} \nabla \cdot \left(v^{(k)} C^{(n-k)} \right) + q_s C_s . \tag{4}$$

Note that the external sink/source term is assumed to be deterministic and can thus be grouped with 0^{th}-order terms and that relaxing this assumption is straightforward [*Zhang*, 2002]. One can separate (4) at different expansion orders with respect to σ_Y:

0^{th} order:

$$\frac{\partial C^{(0)}}{\partial t} = \nabla \cdot \left(D^{(0)} \nabla C^{(0)} \right) - \nabla \cdot \left(v^{(0)} C^{(0)} \right) + q_s C_s , \tag{5a}$$

subject to initial and boundary conditions,

$$C^{(0)}(\boldsymbol{x},0) = C_0(\boldsymbol{x}), \qquad \boldsymbol{x} \in \Omega, \tag{5b}$$

$$C^{(0)}(\boldsymbol{x},t) = C_D(\boldsymbol{x},t), \qquad \boldsymbol{x} \in \Gamma_D, \tag{5c}$$

$$D^{(0)}\nabla C^{(0)} \cdot \boldsymbol{n}(\boldsymbol{x}) = -F, \qquad \boldsymbol{x} \in \Gamma_N. \tag{5d}$$

1^{st}-order:

$$\frac{\partial C^{(1)}}{\partial t} = \nabla \cdot \left(D^{(0)}\nabla C^{(1)}\right) - \nabla \cdot \left(v^{(0)}C^{(1)}\right) + g^{(1)}, \tag{6a}$$

subject to

$$C^{(1)}(\boldsymbol{x},0) = 0, \qquad \boldsymbol{x} \in \Omega, \tag{6b}$$

$$C^{(1)}(\boldsymbol{x},t) = 0, \qquad \boldsymbol{x} \in \Gamma_D, \tag{6c}$$

$$D^{(0)}\nabla C^{(1)} \cdot \boldsymbol{n}(\boldsymbol{x}) = -D^{(1)}\nabla C^{(0)} \cdot \boldsymbol{n}(\boldsymbol{x}), \qquad \boldsymbol{x} \in \Gamma_N, \tag{6d}$$

where

$$g^{(1)} = \nabla \cdot \left(D^{(1)}\nabla C^{(0)} - v^{(1)}C^{(0)}\right). \tag{6e}$$

Similarly one can derive equations at higher expansion orders. Refer to Appendix B for the equations at the 2nd-, 3rd- and general m^{th}-orders, $m \geq 1$. Equations (5) – (6) and (B1) – (B9) are in principle equivalent to the formulations solved by *Hu et al.* [1999] except that our equations have been reformulated into the original form (1) with the stochastic D and v replaced by deterministic $D^{(0)}$ and $v^{(0)}$. The terms involving higher-order $D^{(m)}$ and $v^{(m)}$ at $m \geq 1$ are lumped into the randomness terms $g^{(m)}$.

Unlike that in *Hu et al.* [1999], in the KLME method the higher-order stochastic expansions (6) and (B1) – (B9) are not used to formulate the corresponding moment equations, which can lead to a large dimensionality as mentioned in Introduction. Instead, similar to the KLME flow approach [*Zhang and Lu*, 2004; *Liu et al*., 2006], we further expand $C^{(m)}$, $D^{(m)}$ and $v^{(m)}$ in terms of the orthogonal standard random variables,

$$C^{(m)} = \sum_{i_1,i_2,\ldots,i_m=1}^{\infty} \left(\prod_{j=1}^{m} \xi_{i_j}\right) C^{(m)}_{i_1,i_2,\ldots,i_m}\ ,\ D^{(m)} = \sum_{i_1,i_2,\ldots,i_m=1}^{\infty} \left(\prod_{j=1}^{m} \xi_{i_j}\right) D^{(m)}_{i_1,i_2,\ldots,i_m}\ ,$$

$$v^{(m)} = \sum_{i_1,i_2,\ldots,i_m=1}^{\infty} \left(\prod_{j=1}^{m} \xi_{i_j}\right) v^{(m)}_{i_1,i_2,\ldots,i_m}\ , \tag{7}$$

where $C^{(m)}_{i_1,i_2,\ldots,i_m}$, $D^{(m)}_{i_1,i_2,\ldots,i_m}$ and $v^{(m)}_{i_1,i_2,\ldots,i_m}$ are all deterministic functions; $i_1, i_2, \cdots, i_m$ are referred to as expansion modes at the m^{th} order; ξ_{i_j} are the orthogonal standard random variables. As shown in *Lu and Zhang* [2004] and Appendix A, $D^{(m)}_{i_1,i_2,\ldots,i_m}$ and $v^{(m)}_{i_1,i_2,\ldots,i_m}$ can be calculated after stochastic head solutions are obtained. $C^{(m)}_{i_1,i_2,\ldots,i_m}$ are the quantities to be solved for in this work as they provide the basis to estimate the mean concentrations and concentration variances (as well as other higher concentration moments).

Substituting (7) into (6) and dropping the independent set $\{\xi_i\}$, one obtains the following equations for $C^{(1)}_i$ at the first order mode i,

$$\frac{\partial C^{(1)}_i}{\partial t} = \nabla \cdot \left(D^{(0)} \nabla C^{(1)}_i\right) - \nabla \cdot \left(v^{(0)} C^{(1)}_i\right) + g^{(1)}_i\ , \tag{8a}$$

subject to

$$C^{(1)}_i(\boldsymbol{x},0) = 0, \qquad \boldsymbol{x} \in \Omega\,, \tag{8b}$$

$$C^{(1)}_i(\boldsymbol{x},t) = 0, \qquad \boldsymbol{x} \in \Gamma_D\,, \tag{8c}$$

$$D^{(0)} \nabla C^{(1)}_i \cdot \boldsymbol{n}(\boldsymbol{x}) = -D^{(1)}_i \nabla C^{(0)} \cdot \boldsymbol{n}(\boldsymbol{x}), \qquad \boldsymbol{x} \in \Gamma_N\,, \tag{8d}$$

where

$$g^{(1)}_i = \nabla \cdot \left(D^{(1)}_i \nabla C^{(0)} - v^{(1)}_i C^{(0)}\right). \tag{8e}$$

Similarly one can derive the equations at higher expansion orders and modes. Refer to Appendix B for the equations at the 2nd, 3rd and general m^{th} modes $i_1, i_2, \cdots, i_m$, $m \geq 1$.

Once the deterministic coefficients $C^{(m)}_{i_1,i_2,\ldots,i_m}$ are calculated, one can easily compute the mean concentration and concentration variance by algebraic operations. For example, up to the third order in σ_Y,

$$C \approx \sum_{m=0}^{3} C^{(m)} \ . \tag{9}$$

The mean concentration can be approximated as,

$$< C > \approx \sum_{m=0}^{3} < C^{(m)} > = C^{(0)} + < C^{(2)} > = C^{(0)} + \sum_{i=1}^{\infty} C^{(2)}_{i,i} \ , \tag{10}$$

where the first term in the right-hand side is the zeroth-order mean concentration solution and the second term represents the second-order correction. The terms $C^{(2)}_{i,j}$ for $i \neq j$ disappear because of $< \xi_i \xi_j >= 0$. The first-order correction to the zeroth-order mean is zero due to $< \xi_i >= 0$.

From (9) and (10), one can write the perturbation term up to the third-order in σ_Y as,

$$C' = C - < C > \approx \sum_{m=1}^{3} C^{(m)} - \sum_{i=1}^{\infty} C^{(2)}_{i,i} \ . \tag{11}$$

The concentration variance can be calculated by squaring (11) and then taking the ensemble mean,

$$\sigma_C^2 \approx \sum_{i=1}^{\infty} [C_i^{(1)}]^2 + 2 \sum_{i,j=1}^{\infty} [C_{i,j}^{(2)}]^2 + 6 \sum_{i,j=1}^{\infty} [C_i^{(1)} C_{i,j,j}^{(3)}], \tag{12}$$

where the first term in the right-hand side is the concentration variance up to the first order in the variance of log conductivity σ_Y^2, and the second and third terms represent the second-order corrections in σ_Y^2.

Equations (5), (8), and (B10) – (B18) have the same structure as the original transport equation (1). Therefore, any existing simulator, such as MT3DMS [*Zheng and Wang*, 1999], can

be directly used to solve $C^{(m)}_{i_1,i_2,\ldots,i_m}$. Moreover, due to the same structure of these equations, the left-hand-side coefficient matrix remains unchanged across simulator calculations at different expansion orders and modes, which further allows us to increase the computational efficiency of the KLME approach. Because of the recursive nature, solving (5), (8), and (B10) – (B18) is a sequential process from low to high expansion orders. At the same order, the equations are independent of each other, a feature that enables potential parallel computing for large-scale problems. In this work, $C^{(m)}_{i_1,i_2,\ldots,i_m}$ is evaluated up to the third order. The appropriate solution procedure is: 1) at the current time step, solving the zeroth-order equation (5), 2) solving (8) for $C^{(1)}_i$ at the *1st* order for different modes *i*, 3) continuing the solutions at the *2nd* and *3rd* orders (B10 – B13), 4) computing the mean and variance of concentration using (10) and (12), and 5) adding a time increment and repeating steps 1) – 4), if needed. A new code called "MT3DMS-STO" has been developed to implement numerically the stochastic formulation presented here, in which MT3DMS is employed as a subroutine to calculate $C^{(m)}_{i_1,i_2,\ldots,i_m}$.

3. Illustrative Examples

The validity of the KLME approach for stochastic uncertainty analysis in solute transport modeling is evaluated through a series of 2-D numerical experiments. Results from the proposed methodology are compared to those from the classical MC simulations. The log-transformed random hydraulic conductivity field is assumed to be second-order stationary and follows a separable exponential covariance function,

$$C_Y(\boldsymbol{x}_1,\boldsymbol{x}_2)=\sigma_Y^2\exp[-\frac{|x_1-x_2|}{\eta_x}-\frac{|y_1-y_2|}{\eta_y}], \tag{13}$$

where $\boldsymbol{x}_1 = (x_1, y_1)$ and $\boldsymbol{x}_2 = (x_2, y_2)$ are any two points in the simulation domain, η_x and η_y are the correlation lengths at x and y directions, respectively. In following examples, the correlation lengths are set equal to $\eta_x = \eta_y = 1$ m, and the variance σ_Y^2 varies from 0.25 to 0.5. In the MC simulations 5000 realizations are used. In the KLME approach $C_{i_1}^{(1)}$ are calculated for the first 100 modes, $i_1 = \overline{1,100}$ at the first order; at the second order, $C_{i_1,i_2}^{(2)}$ are calculated for the first 40 by 40 modes, i_1, $i_2 = \overline{1,40}$; and at the third order, $C_{i_1,i_2,i_3}^{(3)}$ are calculated for the first 20 by 20 by 20 modes, i_1, i_2, $i_3 = \overline{1,20}$. Consequently, the total number of mode calculations is 2020. Compared to those in the flow cases [*Zhang and Lu*, 2004; *Liu et al*., 2006], we used a larger number of modes in order to assure the converged statistics for $C_{i_1,i_2,\ldots,i_m}^{(m)}$.

3.1. Numerical Model Setup

The flow and transport domain is 30 m long by 10 m wide (Figure 1). There is no flow across the northern and southern boundaries; on the eastern boundary the hydraulic head is prescribed as constant at 10.0 m; on the western boundary there is a specified influx Q (i.e., 0.01 m^3/d) determined such that an average hydraulic gradient of 0.001 is achieved in the mean flow direction. In the numerical simulations the flux Q is represented by a series of injection wells in the boundary cells. The model is discretized into a block-centered finite difference grid of 121 columns and 40 rows with uniform cell dimensions of 0.25 m by 0.25 m (one extra column added to due to the particular eastern and western boundary configurations). As a result, one correlation length of log conductivity occupies 4 individual cells. The simulated flow field is confined and steady-state.

In the transport simulations, zero concentration gradient is specified across all four boundaries (the dispersive boundary flux F is zero). To explore the effects of local dispersion,

the longitudinal dispersivity α_L is varied from 0.025 m to 1.0 m while the dispersivity ratio between transverse and longitudinal directions (α_{TH}/α_L) remains constant at 0.1. All simulations employ a uniform effective porosity of 0.35, and a molecular diffusion coefficient of 5.0×10^{-4} m^2/d. To reduce the boundary effects, the initial source is placed 4.5 m downstream from the western border (Figure 1). Two different instantaneous source configurations are investigated: (1) point source in which the initial mass is distributed in a single cell and (2) line source in which the initial mass is distributed uniformly in a wide region over 2.5 conductivity correlation lengths (10 cells). Line A – A′ indicates where the simulation results will be analyzed in details.

As in *Liu et al.* [2006], the stochastic expansions for hydraulic heads $h^{(m)}_{i_1,i_2,\ldots,i_m}$ are solved by MODFLOW-2000 [*Harbaugh et al.*, 2000]. For the stochastic expansions of solute concentrations $C^{(m)}_{i_1,i_2,\ldots,i_m}$, as mentioned earlier, MT3DMS is used as the solver. A highly accurate third-order TVD scheme is selected for the advection term to suppress numerical dispersion and artificial oscillations. Other transport terms are solved by an implicit General Conjugate-Gradient (GCG) method. To further ensure numerical accuracy, the transport time step is bounded by a Courant number of 0.50.

3.2. Results and Discussions

3.2.1. $\sigma_Y^2 = 0.25$*, Point Source, Local Disperivities Varied*

In this case we consider a log conductivity variance σ_Y^2 of 0.25, which can be converted to a coefficient of variation of 53.3% for the original conductivity field before log transformation. The type of single-cell point source is examined here. Figure 2 displays the contours of mean concentrations from both MC simulations and the KLME approach at different dimensionless

times. The dimensionless time is defined as $t' = < v > t / \eta_x$, which describes the number of correlation lengths solute plume has traveled in the mean flow direction up to time *t*. The longitudinal dispersivity is set equal to 0.2 m. Figure 2 indicates that compared to the MC solutions, the zeroth-order results, which are based only on the geometric mean conductivity field, overestimate solute spreading at the plume center while show underestimation towards the edges of plume. By adding the second-order correction, i.e., the second term in (12), the results are much improved and the match between the KLME approach and MC simulations becomes significantly better. Due to the anisotropy ratio of 0.1 between the transverse and longitudinal dispersivities, solute spreading is elliptical and elongated along the flow direction. The accuracy of the second-order KLME results to resemble those from MC simulations is consistently good in both transverse and longitudinal directions through different times.

As shown in the previous section, an obvious advantage of the KLME approach is the ease at which concentration variance can be estimated by simple algebraic operations on the mode expansions $C^{(m)}_{i_1,i_2,\ldots,i_m}$. Figure 3 plots the concentration variances calculated by the KLME approach and MC simulations at different times for the same settings as in Figure 2. To facilitate visual observations, only the second-order KLME results (with all three terms in (12)) are provided. It can be seen that despite some local mismatch, the overall concentration variability estimated from the KLME approach are in good agreement with the MC results. Owing to the effects of local dispersion and diffusion, the concentration variances are bimodal in the flow direction, indicating that solute concentrations have the largest variability in the limbs of plume where concentration gradient is high. The bimodal variance behaviors are also observed in *Morales-Casique et al.* [2006b].

Figure 4 presents the mean concentrations and concentration variances along the profile A – A′ in Figures 2 and 3. Being consistent with the contours in Figure 2, the zeroth-order KLME mean concentrations show peaks that are higher than those from MC simulations and underpredict the solute spreading in the areas toward the outside plume edges. The second-order results show significant improvement over the zeroth-order ones and match the solutions from MC simulations very well at both plume center and the outside edges. The accuracy of KLME approach is well preserved through different times. For the concentration variances, the first-order KLME results, which are based on the first term in (12) only, overshoot the bimodal peaks from MC simulations and undershoot the low values at plume center. The second-order corrections, i.e., the second and third terms in (12), help improve the solution accuracy by rectifying both the overestimation at the peaks and underestimation at the middle low values. Compared to the mean concentrations, the second-order improvement on concentration variances is to a lesser degree. This is consistent with the finding in *Morales-Casique et al.* [2006b] that while the mean concentrations can be reproduced adequately under certain circumstances, the errors in concentration variances by the perturbative methods are relatively higher.

Figure 5 displays the standard deviation (σ_{C/C_0}) and coefficient of variation ($CV = \sigma_C / < C >$) of solute concentrations calculated by KLME at the plume center. Results are shown for three different sets of dispersivities and other model settings remain identical to those in Figure 2. Because of the decrease of concentrations, σ_{C/C_0} diminishes continuously with time. The rate of decline becomes smaller at later steps. When dispersion is large, the variability of solute concentrations reduces. Compared to σ_{C/C_0}, CV offers a better indicator of the actual fluctuation of solute concentrations. It is of interest to note that instead of the continuous drop in σ_{C/C_0}, CV becomes more or less stabilized after $t' = 10$ for all three sets of dispersivities

examined. At early times solute migration processes are pre-asymptotic and both σ_{C/C_0} and CV show large variations.

As mentioned earlier, unlike the previous studies [*Kapoor and Gelhar*, 1994a, b; *Zhang and Neuman*, 1996; *Hu et al*., 1999; *Morales-Casique et al.*, 2006a, b] where the local dispersion has been treated as given constants or simple functions of mean velocities, in this work we consider the effects of velocity variations upon local dispersion by including the higher-order $D^{(m)}$ terms in the KLME formulation. Figure 6 compares the results between the case where only zeroth-order dispersion term $D^{(0)}$ is used and the case where the higher-order terms $D^{(1)}$, $D^{(2)}$ and $D^{(3)}$ are also included. The longitudinal dispersivity is 1.0 m and other settings remain identical to those in Figure 2. The mean concentrations did not show much change after including the higher-order dispersion terms and all the results are in excellent agreement with each other. For the concentration variances, however, without considering the contributions of higher-order dispersion terms, the computed values by KLME appear lagged behind consistently. The match on the bimodal peaks also deteriorates if only $D^{(0)}$ is involved. Simulations using other dispersivities indicate that the effects of higher-order dispersion terms on concentration variances become less significant when dispersivities are smaller.

3.2.2. Line Source

To investigate the effects of different source configurations on the performance of the KLME approach, here we consider the line source and maintain other model settings the same as those in Figure 2. Figure 7 shows the results under the new source scenario. Compared to the previous single-cell point source, both the mean concentrations and concentration variances are larger in the line source due to more mass initially in the system. Nonetheless, similar observations can be made regarding the match between the KLME approach and MC simulations.

For the mean concentrations, the zeroth-order KLME results overshoot the high peak from MC at the plume center and undershoot the low values at the plume edges, and the second-order correction term is able to improve significantly the agreement in both areas. For concentration variances, the first-order KLME results overestimate at the bimodal peaks and underestimate at the middle low values and the second-order correction terms are capable of correcting both to a high degree. Despite the different source configurations, the coefficient of variation calculated at plume center is smaller for the line source than that for the point source, indicating that the prediction uncertainty gets reduced as the plume covers a larger portion of the flow field.

3.2.3. $\sigma_Y^2 = 0.5$

In this section we explore the performance of the KLME approach when the conductivity variance σ_Y^2 is increased up to 0.5. The corresponding coefficient of variation is 80.5% for the original conductivity field. Figure 8 displays the mean concentrations and concentration variances calculated by the KLME approach and MC simulations at $\sigma_Y^2 = 0.5$. Other settings remain identical to those in Figure 2. It is seen that when the conductivity variance increases to 0.5, the second-order KLME results are still accurate in matching the mean concentrations from MC simulations. For the concentration variances, however, the match between the KLME approach and MC simulations becomes worse especially at the plume center. The second-order KLME corrections overpredict largely the middle low values from MC simulations. The constraint of conductivity variance on solution accuracy has also been observed in *Morales-Casique et al.* [2006b] where the iterative perturbative solutions of advective-dispersive transport problems are found only applicable for $\sigma_Y^2 < 0.3$. As mentioned in Introduction, this limitation may be conceivably overcome by the new multiscale decomposition approach proposed in *Neuman* [2006].

3.2.4. Pumping Well Included

To explore the performance of the KLME approach under the influence of external stresses, a pumping well is added and placed at the center of domain at a rate of 0.02 m^3/d. An observation well is also put in 5 m upstream of the pumping well. The zeroth-order concentrations at $t' = 3$ for the case in Figure 2 are applied as the initial distributed source. Other model settings are kept the same as those in Figure 2. Figure 9 shows the mean concentrations and concentration variances observed at the pumping and observation wells at different times. For the mean concentrations at the two wells, both the zeroth- and second-order KLME results are close to those determined from MC simulations, and the second-order correction yields some modest improvement. For concentration variances, the breakthrough curves are bimodal, the first-order results overshoot at the peaks and underpredict at the middle low values, and the second-order corrections are able to improve the match between KLME and MC simulations to a certain degree.

At the pumping well the mean concentration peak arrives at approximately 1700 days after the start of model simulation. It can be seen that the largest variability of solute concentrations at the pumping well does not correspond to the breakthrough concentration peak. The first variance peak appears shortly before the mean concentration peak arrival (~ 1600 days), and the second shows up in later times after the main plume passes by (~ 3500 days). This is because the concentration fluctuations, as shown above, occur most significantly at the limbs instead of the center area on the spatial solute distribution.

4. Summary and Conclusions

In this paper we have extended the Karhunen-Loeve based moment equation (KLME) approach proposed by *Zhang and Lu* [2004] to solving subsurface solute transport problems in randomly heterogeneous media. The KLME approach is based on an innovative combination of Karhunen-Loeve (KL) decomposition of the underlying random conductivity field and the perturbative and polynomial expansions of other dependent variables including the hydraulic heads, flow velocities, dispersion terms and solute concentrations. The equations obtained in this approach are recursive and can be solved sequentially from low to high orders. The structure of these equations has been formulated in the same form as the original governing equations such that any existing simulator, such as MT3DMS [*Zheng and Wang*, 1999], can be directly applied as the solver. The theoretical derivations presented in this work have been numerically implemented in a code called "MT3DMS-STO".

The validity of the KLME approach has been evaluated against the classical Monte Carlo (MC) simulations in a series of 2-D numerical experiments under different flow and transport conditions. In all different cases examined, the KLME approach yields a good representation of the mean concentrations from MC simulations after the second-order correction. For concentration variances, despite some local mismatch, the KLME approach is effective in describing the overall trend at the conductivity variance σ_Y^2=0.25. When σ_Y^2 is increased to 0.5, the differences on concentration variances between the KLME approach and MC simulations become substantial, suggesting a need for the third- and even higher-order corrections.

Some specific conclusions can be drawn from the results of the 2-D examples in this work. First, the zeroth-order mean concentrations tend to overpredict at plume center and underpredict at the outside plume edges. The second-order KLME correction is able to rectify the mismatch in both areas. The agreement between the KLME approach and MC simulations is

consistently good through different times. Second, the first-order KLME concentration variances tend to overestimate the bimodal peaks calculated by MC at the limbs of plume where the concentration gradient is large. At plume center, the concentration variances are small and the first-order KLME results show underestimation. Similar to the mean concentrations, the second-order correction terms are able to improve the variance mismatch at both the peaks and middle low values, but to a lesser degree. Third, when local dispersivities are large, the variability of concentrations becomes small. Neglecting the higher-order dispersion terms, as commonly done in previous studies [*Kapoor and Gelhar*, 1994a, b; *Zhang and Neuman*, 1996; *Hu et al.*, 1999; *Morales-Casique et al.*, 2006a, b], does not materially affect the mean concentration results but generates a large error on concentration variances. Fourth, in the line source case where more mass is initially distributed into the system, although the absolute magnitudes of mean concentrations and concentration variances increase, the actual variability of solute concentrations, measured by the *CV*, decreases as the initial plume size increases. Fifth, when the variance of random conductivity field increases from 0.25 to 0.5, the KLME approach is still accurate in representing the mean concentrations in MC. However, the mismatch on the concentration variances becomes more pronounced, especially at the middle low values where the second-order correction terms overpredict significantly. Similar constraint has also been observed in *Morales-Casique et al.* [2006b] where perturbative solutions of advective-dispersive transport problems are found applicable for σ_Y^2 <0.3. This limitation may be overcome by the new multiscale decomposition approach proposed in *Neuman* [2006]. Finally, the validity of KLME approach maintains under the influence of a pumping well. The large concentration variances appear before and after the peak arrival on the mean concentration breakthrough curve at the pumping well.

Acknowledgments This work is partially supported by Los Alamos National Laboratory through its LDRD/DR project "Science of Geological Carbon Sequestration", by U.S. National Science Foundation through its grant OISE#0511496, by American Chemical Society through its Petroleum Research Fund (Grant #42109-AC9), and by Chinese Natural Science Foundation through its grant 50310444.

Appendix A: Higher-order Expansions of Dispersion Tensor *D*

Let $W = 1/|v|$, $U_x = v_x{}^2$, $U_y = v_y{}^2$, $U_z = v_z{}^2$, $D_{xxL} = \alpha_L v_x{}^2/|v|$, $D_{xxTH} = \alpha_{TH} v_y{}^2/|v|$ and $D_{xxTV} = \alpha_{TV} v_z{}^2/|v|$, then

$$D_{xxL} = \alpha_L U_x W,\ D_{xxTH} = \alpha_{TH} U_y W,\ D_{xxTV} = \alpha_{TV} U_z W, \tag{A1}$$

and

$$D_{xx} = D_{xxL} + D_{xxTH} + D_{xxTV} + D^*. \tag{A2}$$

By approximating the pore water velocities up to the third order in σ_Y, W can be expanded as,

$$W = 1/|v| = (v_x^2 + v_y^2 + v_z^2)^{-1/2}$$

$$\approx \left((v_x^{(0)} + v_x^{(1)} + v_x^{(2)} + v_x^{(3)})^2 + (v_y^{(0)} + v_y^{(1)} + v_y^{(2)} + v_y^{(3)})^2 + (v_z^{(0)} + v_z^{(1)} + v_z^{(2)} + v_z^{(3)})^2\right)^{-1/2}$$

$$= \frac{1}{|v^{(0)}|}(1+T)^{-1/2}, \tag{A3}$$

where

$$|v^{(0)}| = (v_x^{(0)} + v_y^{(0)} + v_z^{(0)})^{1/2}, \tag{A4}$$

and

$$T = \frac{1}{|v^{(0)}|^2}\Big(2(v_x^{(0)}v_x^{(1)} + v_y^{(0)}v_y^{(1)} + v_z^{(0)}v_z^{(1)}) + (v_x^{(1)2} + v_y^{(1)2} + v_z^{(1)2}) +$$

$$2(v_x^{(0)}v_x^{(2)} + v_y^{(0)}v_y^{(2)} + v_z^{(0)}v_z^{(2)}) + 2(v_x^{(0)}v_x^{(3)} + v_y^{(0)}v_y^{(3)} + v_z^{(0)}v_z^{(3)}) +$$

$$2(v_x^{(1)}v_x^{(2)} + v_y^{(1)}v_y^{(2)} + v_z^{(1)}v_z^{(2)}) + O(\sigma_Y^4)\Big). \tag{A5}$$

Applying Taylor expansion to (A3) and truncating it at the third order give,

$$W \approx \frac{1}{|v^{(0)}|}\left(1 - \frac{1}{2}T + \frac{3}{8}T^2 - \frac{5}{16}T^3\right). \tag{A6}$$

Substituting (A5) in (A6) and separating W at different orders $W^{(m)}$ yield,

$$W^{(0)} = \frac{1}{|v^{(0)}|}, \tag{A7}$$

$$W^{(1)} = \frac{-1}{|v^{(0)}|^3}\left(v_x^{(0)}v_x^{(1)} + v_y^{(0)}v_y^{(1)} + v_z^{(0)}v_z^{(1)}\right), \tag{A8}$$

$$W^{(2)} = \frac{-1}{2|v^{(0)}|^3}(v_x^{(1)^2} + v_y^{(1)^2} + v_z^{(1)^2} + 2v_x^{(0)}v_x^{(2)} + 2v_y^{(0)}v_y^{(2)} + 2v_z^{(0)}v_z^{(2)}) +$$

$$\frac{3}{2|v^{(0)}|^5}(v_x^{(0)}v_x^{(1)} + v_y^{(0)}v_y^{(1)} + v_z^{(0)}v_z^{(1)})^2, \tag{A9}$$

$$W^{(3)} = \frac{-1}{|v^{(0)}|^3}(v_x^{(0)}v_x^{(3)} + v_y^{(0)}v_y^{(3)} + v_z^{(0)}v_z^{(3)} + v_x^{(1)}v_x^{(2)} + v_y^{(1)}v_y^{(2)} + v_z^{(1)}v_z^{(2)}) +$$

$$\frac{3}{2|v^{(0)}|^5}(v_x^{(0)}v_x^{(1)} + v_y^{(0)}v_y^{(1)} + v_z^{(0)}v_z^{(1)})(v_x^{(1)^2} + v_y^{(1)^2} + v_z^{(1)^2} + 2v_x^{(0)}v_x^{(2)} + 2v_y^{(0)}v_y^{(2)} + 2v_z^{(0)}v_z^{(2)})$$

$$-\frac{5}{2|v^{(0)}|^7}(v_x^{(0)}v_x^{(1)} + v_y^{(0)}v_y^{(1)} + v_z^{(0)}v_z^{(1)})^3. \tag{A10}$$

Up to third order in σ_Y, $U_x{}^2$ can be approximated as,

$$U_x \approx (v_x^{(0)} + v_x^{(1)} + v_x^{(2)} + v_x^{(3)})^2. \tag{A11}$$

Separating U_x at different orders gives,

$$U_x^{(0)} = (v_x^{(0)})^2, \tag{A12}$$

$$U_x^{(1)} = 2v_x^{(0)}v_x^{(1)}, \tag{A13}$$

$$U_x^{(2)} = 2v_x^{(0)}v_x^{(2)} + (v_x^{(1)})^2, \tag{A14}$$

$$U_x^{(3)} = 2v_x^{(0)}v_x^{(3)} + 2v_x^{(1)}v_x^{(2)}. \tag{A15}$$

D_{xxL} can be approximated as,

$$D_{xxL} = \alpha_L U_x W \approx \alpha_L (U_x^{(0)} + U_x^{(1)} + U_x^{(2)} + U_x^{(3)})(W^{(0)} + W^{(1)} + W^{(2)} + W^{(3)}). \quad \text{(A16)}$$

Separating D_{xxL} at different orders gives,

$$D_{xxL}^{(0)} = \alpha_L U_x^{(0)} W^{(0)}, \quad \text{(A17)}$$

$$D_{xxL}^{(1)} = \alpha_L (U_x^{(0)} W^{(1)} + U_x^{(1)} W^{(0)}), \quad \text{(A18)}$$

$$D_{xxL}^{(2)} = \alpha_L (U_x^{(0)} W^{(2)} + U_x^{(1)} W^{(1)} + U_x^{(2)} W^{(0)}), \quad \text{(A19)}$$

$$D_{xxL}^{(3)} = \alpha_L (U_x^{(0)} W^{(3)} + U_x^{(1)} W^{(2)} + U_x^{(2)} W^{(1)} + U_x^{(3)} W^{(0)}). \quad \text{(A20)}$$

Similarly one can formulate the expansions of D_{xxTH} and D_{xxTV} at different orders. D_{xx} (equation A2) can then be approximated up to the third order in σ_Y as,

$$D_{xx} \approx D_{xx}^{(0)} + D_{xx}^{(1)} + D_{xx}^{(2)} + D_{xx}^{(3)}, \quad \text{(A21)}$$

where

$$D_{xx}^{(0)} = D_{xxL}^{(0)} + D_{xxTH}^{(0)} + D_{xxTV}^{(0)} + D^*, \quad \text{(A22)}$$

$$D_{xx}^{(1)} = D_{xxL}^{(1)} + D_{xxTH}^{(1)} + D_{xxTV}^{(1)}, \quad \text{(A23)}$$

$$D_{xx}^{(2)} = D_{xxL}^{(2)} + D_{xxTH}^{(2)} + D_{xxTV}^{(2)}, \quad \text{(A24)}$$

$$D_{xx}^{(3)} = D_{xxL}^{(3)} + D_{xxTH}^{(3)} + D_{xxTV}^{(3)}. \quad \text{(A25)}$$

Because of the identical structure as shown in (3), the above procedure applies exactly to the derivations of D_{yy} and D_{zz} and will be not repeated here.

To derive the expansions for the cross terms $D_{xy} = D_{yx} = (\alpha_L - \alpha_{TH}) v_x v_y / |v|$, letting $R_{xy} = v_x v_y$ and substituting in the velocity expansions up to the third order, and then separating it at different orders, one obtains,

$$R_{xy}^{(0)} = v_x^{(0)} v_y^{(0)}, \quad \text{(A26)}$$

$$R_{xy}^{(1)} = v_x^{(0)} v_y^{(1)} + v_x^{(1)} v_y^{(0)}, \tag{A27}$$

$$R_{xy}^{(2)} = v_x^{(0)} v_y^{(2)} + v_x^{(1)} v_y^{(1)} + v_x^{(2)} v_y^{(0)}, \tag{A28}$$

$$R_{xy}^{(3)} = v_x^{(0)} v_y^{(3)} + v_x^{(1)} v_y^{(2)} + v_x^{(2)} v_y^{(1)} + v_x^{(3)} v_y^{(0)}. \tag{A29}$$

$D_{xy} = D_{yx}$ can then be approximated up to the third order in σ_Y as,

$$D_{xy} = D_{yx} \approx D_{xy}^{(0)} + D_{xy}^{(1)} + D_{xy}^{(2)} + D_{xy}^{(3)}, \tag{A30}$$

where

$$D_{xy}^{(0)} = D_{yx}^{(0)} = (\alpha_L - \alpha_{TH}) R_{xy}^{(0)} W^{(0)}, \tag{A31}$$

$$D_{xy}^{(1)} = D_{yx}^{(1)} = (\alpha_L - \alpha_{TH})(R_{xy}^{(0)} W^{(1)} + R_{xy}^{(1)} W^{(0)}), \tag{A32}$$

$$D_{xy}^{(2)} = D_{yx}^{(2)} = (\alpha_L - \alpha_{TH})(R_{xy}^{(0)} W^{(2)} + R_{xy}^{(1)} W^{(1)} + R_{xy}^{(2)} W^{(0)}), \tag{A33}$$

$$D_{xy}^{(3)} = D_{yx}^{(3)} = (\alpha_L - \alpha_{TH})(R_{xy}^{(0)} W^{(3)} + R_{xy}^{(1)} W^{(2)} + R_{xy}^{(2)} W^{(1)} + R_{xy}^{(3)} W^{(0)}). \tag{A34}$$

The derivations of $D_{xz} = D_{zx}$ and $D_{yz} = D_{zy}$ are similar and thus not repeated here. The expansion mode coefficients for the entire dispersion tensor $D_{i_1,i_2,\ldots,i_m}^{(m)}$ can be directly obtained by substituting in the respective velocity expansion coefficients $v_{i_1,i_2,\ldots,i_m}^{(m)}$ [*Lu and Zhang*, 2004].

Appendix B: Derivations of Higher-order Equations

Separating (4) at the 2nd and 3rd orders, one obtains the following equations for concentration expansions $C^{(2)}$ and $C^{(3)}$:

$$\frac{\partial C^{(2)}}{\partial t} = \nabla \cdot \left(D^{(0)} \nabla C^{(2)} \right) - \nabla \cdot \left(v^{(0)} C^{(2)} \right) + g^{(2)}, \tag{B1}$$

$$\frac{\partial C^{(3)}}{\partial t} = \nabla \cdot \left(D^{(0)} \nabla C^{(3)} \right) - \nabla \cdot \left(v^{(0)} C^{(3)} \right) + g^{(3)}, \tag{B2}$$

where

$$g^{(2)} = \nabla \cdot \left(D^{(2)} \nabla C^{(0)} + D^{(1)} \nabla C^{(1)} - v^{(2)} C^{(0)} - v^{(1)} C^{(1)} \right). \tag{B3}$$

$$g^{(3)} = \nabla \cdot \left(D^{(3)} \nabla C^{(0)} + D^{(2)} \nabla C^{(1)} + D^{(1)} \nabla C^{(2)} - v^{(3)} C^{(0)} - v^{(2)} C^{(1)} - v^{(1)} C^{(2)} \right). \tag{B4}$$

In general, at m^{th}-order, $m \geq 1$:

$$\frac{\partial C^{(m)}}{\partial t} = \nabla \cdot \left(D^{(0)} \nabla C^{(m)} \right) - \nabla \cdot \left(v^{(0)} C^{(m)} \right) + g^{(m)}, \tag{B5}$$

subject to initial and boundary conditions,

$$C^{(m)}(\boldsymbol{x}, 0) = 0, \qquad \boldsymbol{x} \in \Omega, \tag{B6}$$

$$C^{(m)}(\boldsymbol{x}, t) = 0, \qquad \boldsymbol{x} \in \Gamma_D, \tag{B7}$$

$$D^{(0)} \nabla C^{(m)} \cdot \boldsymbol{n}(\boldsymbol{x}) = - \sum_{i=0}^{m-1} D^{(m-i)} \nabla C^{(i)} \cdot \boldsymbol{n}(\boldsymbol{x}), \qquad \boldsymbol{x} \in \Gamma_N, \tag{B8}$$

where

$$g^{(m)} = \nabla \cdot \sum_{i=0}^{m-1} (D^{(m-i)} \nabla C^{(i)} - v^{(m-i)} C^{(i)}). \tag{B9}$$

The governing equations for the 2nd- and 3rd-order mode expansion coefficients $C_{i,j}^{(2)}$ and $C_{i,j,k}^{(3)}$ are,

$$\frac{\partial C_{i,j}^{(2)}}{\partial t} = \nabla \cdot \left(D^{(0)} \nabla C_{i,j}^{(2)} \right) - \nabla \cdot \left(v^{(0)} C_{i,j}^{(2)} \right) + g_{i,j}^{(2)}, \tag{B10}$$

$$\frac{\partial C_{i,j,k}^{(3)}}{\partial t} = \nabla \cdot \left(D^{(0)} \nabla C_{i,j,k}^{(3)} \right) - \nabla \cdot \left(v^{(0)} C_{i,j,k}^{(3)} \right) + g_{i,j,k}^{(3)}, \tag{B11}$$

where

$$g_{i,j}^{(2)} = \nabla \cdot \left(D_{i,j}^{(2)} \nabla C^{(0)} + (D_i^{(1)} \nabla C_j^{(1)} + D_j^{(1)} \nabla C_i^{(1)})/2 - v_{i,j}^{(2)} C^{(0)} - (v_i^{(1)} C_j^{(1)} + v_j^{(1)} C_i^{(1)})/2 \right). \tag{B12}$$

$$g_{i,j,k}^{(3)} = \nabla \cdot \left(D_{i,j,k}^{(3)} \nabla C^{(0)} + (D_{i,j}^{(2)} \nabla C_k^{(1)} + D_{i,k}^{(2)} \nabla C_j^{(1)} + D_{j,k}^{(2)} \nabla C_i^{(1)})/3 + \right.$$

$$(D_i^{(1)}\nabla C_{j,k}^{(2)} + D_j^{(1)}\nabla C_{i,k}^{(2)} + D_k^{(1)}\nabla C_{i,j}^{(2)})/3 - v_{i,j,k}^{(3)}C^{(0)} -$$

$$(v_{i,j}^{(2)}C_k^{(1)} + v_{i,k}^{(2)}C_j^{(1)} + v_{j,k}^{(2)}C_i^{(1)})/3 - (v_i^{(1)}C_{j,k}^{(2)} + v_j^{(1)}C_{i,k}^{(2)} + v_k^{(1)}C_{i,j}^{(2)})/3\Big). \quad \text{(B13)}$$

In general, at m^{th}-order and mode $i_1, i_2, ..., i_m$, $m \geq 1$:

$$\frac{\partial C_{i_1,i_2,...,i_m}^{(m)}}{\partial t} = \nabla\cdot\left(D^{(0)}\nabla C_{i_1,i_2,...,i_m}^{(m)}\right) - \nabla\cdot\left(v^{(0)}C_{i_1,i_2,...,i_m}^{(m)}\right) + g_{i_1,i_2,...,i_m}^{(m)}, \quad \text{(B14)}$$

subject to initial and boundary conditions,

$$C_{i_1,i_2,...,i_m}^{(m)}(\boldsymbol{x},0) = 0, \qquad \boldsymbol{x}\in\Omega, \quad \text{(B15)}$$

$$C_{i_1,i_2,...,i_m}^{(m)}(\boldsymbol{x},t) = 0, \qquad \boldsymbol{x}\in\Gamma_D, \quad \text{(B16)}$$

$$D^{(0)}\nabla C_{i_1,i_2,...,i_m}^{(m)}\cdot\boldsymbol{n}(\boldsymbol{x}) = -\sum_{k=0}^{m-1}\left[\frac{(m-k)!}{m!}\sum_{P_{i_1,i_2,...,i_m}}(D_{i_{k+1},...,i_m}^{(m-k)}\nabla C_{i_1,...,i_k}^{(k)})\right]\cdot\boldsymbol{n}(\boldsymbol{x}), \quad \boldsymbol{x}\in\Gamma_N, \quad \text{(B17)}$$

where

$$g_{i_1,i_2,...,i_m}^{(m)} = \nabla\cdot\sum_{k=0}^{m-1}\left[\frac{(m-k)!}{m!}\sum_{P_{i_1,i_2,...,i_m}}(D_{i_{k+1},...,i_m}^{(m-k)}\nabla C_{i_1,...,i_k}^{(k)} - v_{i_{k+1},...,i_m}^{(m-k)}C_{i_1,...,i_k}^{(k)})\right]. \quad \text{(B18)}$$

The summation $\sum_{P_{i_1,i_2,...,i_m}}$ in (B17) and (B18) is taken over a subset of the permutation of $\{i_1, i_2, ..., i_m\}$ in which repeated terms of $D_{i_{k+1},...,i_m}^{(m-k)}\nabla C_{i_1,...,i_k}^{(k)}$ and $v_{i_{k+1},...,i_m}^{(m-k)}C_{i_1,...,i_k}^{(k)}$ are excluded. For example, $\sum_{P_{i,j,k}} D_i^{(1)}\nabla C_{j,k}^{(2)} = D_i^{(1)}\nabla C_{j,k}^{(2)} + D_j^{(1)}\nabla C_{i,k}^{(2)} + D_k^{(1)}\nabla C_{i,j}^{(2)}$. $D_i^{(1)}\nabla C_{k,j}^{(2)}$ is identical to $D_i^{(1)}\nabla C_{j,k}^{(2)}$ and thereby excluded as $C_{j,k}^{(2)}$ calculated this way is symmetric with respect to its subscript indices. Furthermore, due to the symmetry, we only need to solve $C_{i_1,i_2,...,i_m}^{(m)}$ for $i_1 \leq i_2 \leq \cdots \leq i_m$ as $C_{i_1,i_2,...,i_m}^{(m)}$ for $i_1 > i_2 > \cdots > i_m$ can be directly obtained by simple manipulation of the subscripts.

References

Burnett, R. D. and E. O. Frind (1987), An alternating direction Galerkin technique for simulation of groundwater contaminant transport in three dimensions: 2. Dimensionality effects, *Water Resour. Res.*, 23(4), 695–705.

Chen, M., D. Zhang, A. A. Keller, and Z. Lu (2005), A stochastic analysis of steady state two-phase flow in heterogeneous media, *Water Resour. Res.*, 41, W01006, doi:10.1029/2004WR003412.

Cushman, J.H. (1997), *The Physics of Fluids in Hierarchical Porous Media: Angstroms to Miles*, Kluwer Academic Publishers, Norwell, MA.

Dagan, G. (1984), Solute transport in heterogeneous formations, *J. Fluid Mech.*, 145, 151–177.

Dagan, G. (1989), *Flow and Transport in Porous Formations*, Springer, New York.

Dagan, G. and A. Fiori (1997), The influence of pore-scale dispersion on concentration statistical moments in transport through heterogeneous aquifers, *Water Resour. Res.*, 33(7), 1595–1605.

Dagan, G., V. Cvetkovic and A. Shapiro (1992), A Solute Flux Approach to Transport in Heterogeneous Formations 1. The General Framework, *Water Resour. Res.*, 28(5), 1369–1376.

Gelhar, L. W., A. L. Gutiahr, and R. L. Naff (1979), Stochastic analysis of macrodispersion in a stratified aquifer, *Water Resour. Res.*, 15(6), 1387–1397.

Gelhar, L.W. (1993), *Stochastic Subsurface Hydrology*, Prentice-Hall, Englewood Cliffs, NJ.

Ghanem, R. and S. Dham (1998), Stochastic finite element analysis for multiphase flow in heterogeneous porous media, *Transp. Porous Media*, 32, 239– 262.

Ghanem, R. and D. Spanos (1991), *Stochastic Finite Elements: A Spectral Approach,* Springer, New York.

Graham, W.D. and D. McLaughlin (1989), Stochastic analysis of nonstationary subsurface solute transport, 1. Unconditional moments, *Water Resour. Res.*, 25, 215-232.

Harbaugh, A.W., E.R. Banta, M.C. Hill and M.G. McDonald, *MODFLOW-2000, the U.S. Geological Survey Modular Ground-Water Model – User Guide to Modularization Concepts and the Ground-Water Flow Processes*, U.S. Geological Survey Open-File Report 00-92, 121p.

Hassan, A. E., J. H. Cushman and J. W. Delleur (1998), A Monte Carlo assessment of Eulerian flow and transport perturbation models, *Water Resour. Res*., *34*, 1143–1163.

Hu, B. X., A. E. Hassan and J. H. Cushman (1999), Eulerian solutions of O(σ_v^N) for the stochastic transport problem for conservative tracers coupled with O(σ_f^4) solutions for the flow problem in an infinite domain, *Water Resour. Res.*, 35(12), 3685–3697.

Kapoor, V. and L. W. Gelhar (1994a), Transport in three-dimensional heterogeneous aquifers, 1, Dynamics of concentration fluctuations, *Water Resour. Res.*, 30(6), 1775–1788.

Kapoor, V. and L. W. Gelhar (1994b), Transport in three-dimensional heterogeneous aquifers, 2, Predictions and observations of concentration fluctuations, *Water Resour. Res.*, 30(6), 1789–1801.

Liu, G., D. Zhang and Z. Lu (2006), Stochastic Uncertainty Analysis for Unconfined Flow Systems, *Water Resour. Res.*, in press.

Lu, Z. and D. Zhang (2004), Conditional Simulations of Flow in Randomly Heterogeneous Porous Media using a KL-based Moment-equation Approach, *Advances in Water Resour.*, 27:859-874, 2004.

Morales-Casique, E., S. P. Neuman, and A. Guadagnini (2006a), Nonlocal and localized analyses of nonreactive solute transport in bounded randomly heterogeneous porous media: Theoretical framework, *Adv. Water Resour.*, in press.

Morales-Casique, E., S. P. Neuman, and A. Guadagnini (2006b), Nonlocal and localized analyses of nonreactive solute transport in bounded randomly heterogeneous porous media: Computational analysis, *Adv. Water Resour.*, in press.

Neuman, S. P. (1993), Eulerian-Lagrangian theory of transport in space-time nonstationary velocity fields: Exact nonlocal formalism by conditional moments and weak approximations, *Water Resour. Res.*, 29(3), 633–645.

Neuman, S.P. (2006), Blueprint for perturbative solution of flow and transport in strongly heterogeneous composite media using fractal and variational multiscale decomposition. *Water Resour Res*., 42, W06D04, doi:10.1029/2005WR004315.

Neuman, S. P. and Y. K. Zhang (1990), A quasi-linear theory of non-Fickian and Fickian subsurface dispersion: 1. Theoretical analysis with application to isotropic media, *Water Resour. Res.*, 26(5), 887–902.

Rubin, Y. (2003), *Applied stochastic hydrogeology*, Oxford University Press, Oxford.

Yang, J., D. Zhang, and Z. Lu (2004), Stochastic analysis of saturated-unsaturated flow in heterogeneous media by combining Karhunen-Loeve expansion and perturbation method, *J. Hydrol*., 294, 18– 38.

Zhang, D. (2002), *Stochastic Methods for Flow in Porous Media: Coping With Uncertainties*, 350 pp., Elsevier, New York.

Zhang, D., and S. P. Neuman (1995), Eulerian-Lagrangian analysis of transport conditioned on hydraulic data, 1, Analytical-numerical approach, *Water Resour. Res.*, 31(1), 39–51.

Zhang, D. and S. P. Neuman (1996), Effect of local dispersion on solute transport in randomly heterogeneous media, *Water Resour. Res.*, 32(9), 2715–2723.

Zhang, D. and Z. Lu (2004), An efficient, high-order perturbation approach for flow in random porous media via Karhunen–Loéve and polynomial expansions, *J. Comput. Phys.*, 194, 773–794.

Zheng, C., and P. P. Wang (1999), MT3DMS: A modular three-dimensional multispecies model for simulation of advection, dispersion and chemical reactions of contaminants in groundwater systems; documentation and user's guide, *Contract Rep. SERDP-99-1*, U.S. Army Eng. Res. and Dev. Cent., Vicksburg, Miss.

LIST OF FIGURES

Figure 1. Schematic diagram of model setup in the 2-D examples. Line A-A′ indicates the location along which results are compared between the KLME and MC approaches.

Figure 2. Normalized mean concentrations ($< C/C_0 >$) calculated from MC simulations and the KLME approach at three different times: **(a)** $t' = 6$, **(b)** $t' = 12$, **(c)** $t' = 18$. C_0 is initial source concentration. Fill contours stand for the results from MC simulations, dash lines for the zeroth-order solutions and solid lines for the results after the second-order correction in KLME. For reference, the initial source is also shown on (**a**).

Figure 3. Concentration variances (σ^2_{C/C_0}) calculated from MC simulations and the KLME approach at three different times: **(a)** $t' = 6$, **(b)** $t' = 12$, **(c)** $t' = 18$. To facilitate visual comparisons, only the second-order results are shown in the KLME approach.

Figure 4. The means and variances of concentrations along the profile A-A' at three different times in Figure 2: **(a)** $t' = 6$, **(b)** $t' = 12$, **(c)** $t' = 18$. The legends for **(b)** and **(c)** are provided in **(a)**.

Figure 5. The standard deviation (σ_{C/C_0}) and coefficient of variation (*CV*) at plume centers calculated in the KLME approach at different dispersivities.

Figure 6. Comparisons of the mean concentrations and concentration variances calculated by the KLME approach between the case where only zeroth-order dispersion term $D^{(0)}$ is used and the case where the higher-order terms $D^{(1)}$, $D^{(2)}$ and $D^{(3)}$ are also included. Results are shown for $t' = 12$.

Figure 7. Calculated results in the line source: (**a**) $< C/C_0 >$ and (**b**) σ^2_{C/C_0} along the profile A-A' at $t' = 12$, and (**c**) *CV* at plume center through different times. The legend for (**b**) is provided in (**a**). The *CV* calculated for the point source is also shown for comparison.

Figure 8. The means and variances of concentrations along the profile A-A' at $t' = 12$ when σ^2_Y is increased to 0.5.

Figure 9. The breakthrough curves for the mean concentrations and concentration variances at (**a**) the pumping well and (**b**) the observation well. The legend for (**b**) is shown in (**a**).

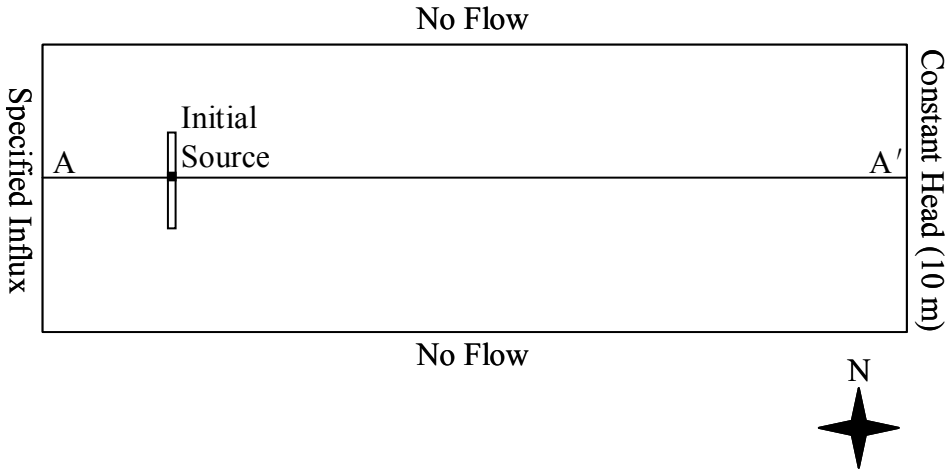


Figure 1. Schematic diagram of model setup in the 2-D examples. Line A-A′ indicates where the simulation results are compared between the KLME and MC approaches in details.

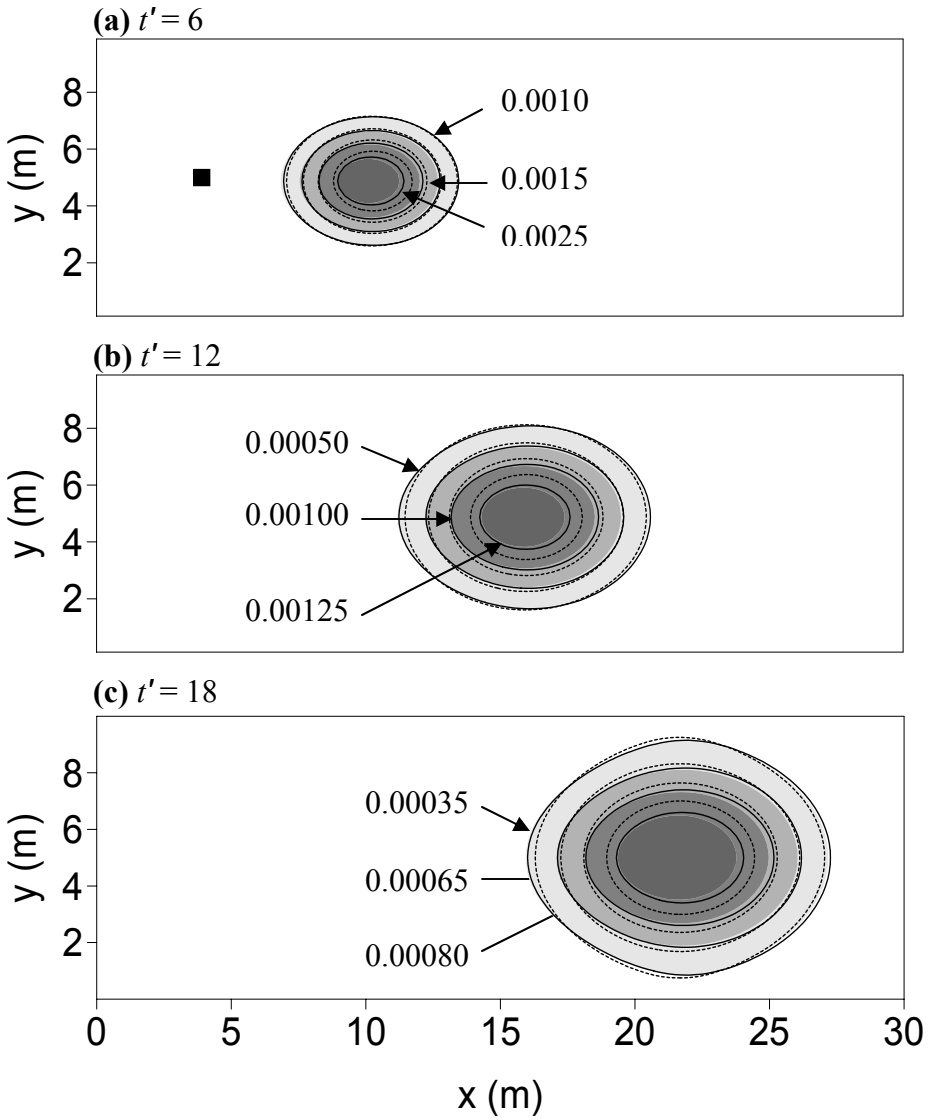


Figure 2. Normalized mean concentrations ($< C / C_0 >$) calculated from MC simulations and the KLME approach at three different times: **(a)** $t' = 6$, **(b)** $t' = 12$, **(c)** $t' = 18$. C_0 is the initial source concentration. Filled contours stand for the results from MC simulations, dash lines for the zeroth-order solutions and solid lines for the results after the second-order correction in KLME. For reference, the initial source is also shown on (**a**).

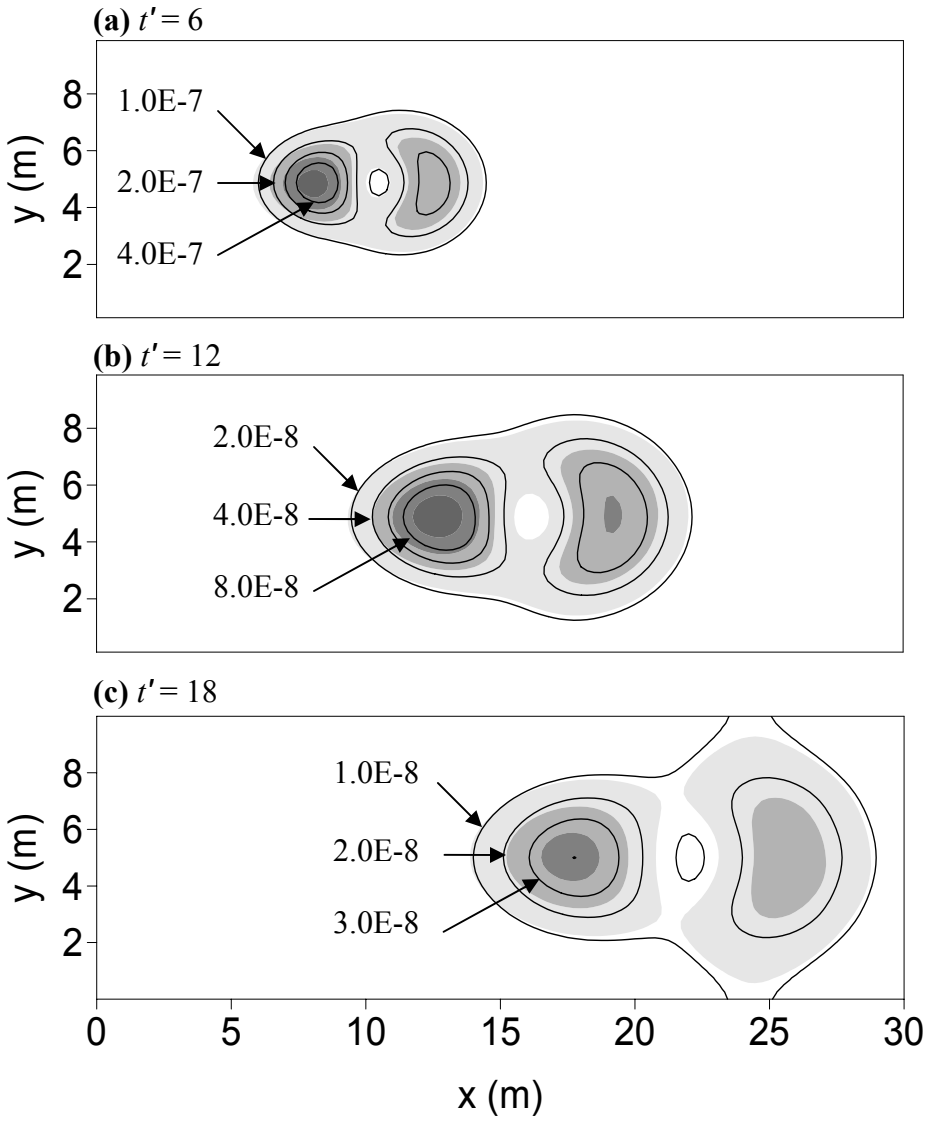


Figure 3. Concentration variances (σ^2_{C/C_0}) calculated from MC simulations and the KLME approach at three different times: **(a)** t' = 6, **(b)** t' = 12, **(c)** t' = 18. To facilitate visual comparisons, only the second-order results are shown in the KLME approach.

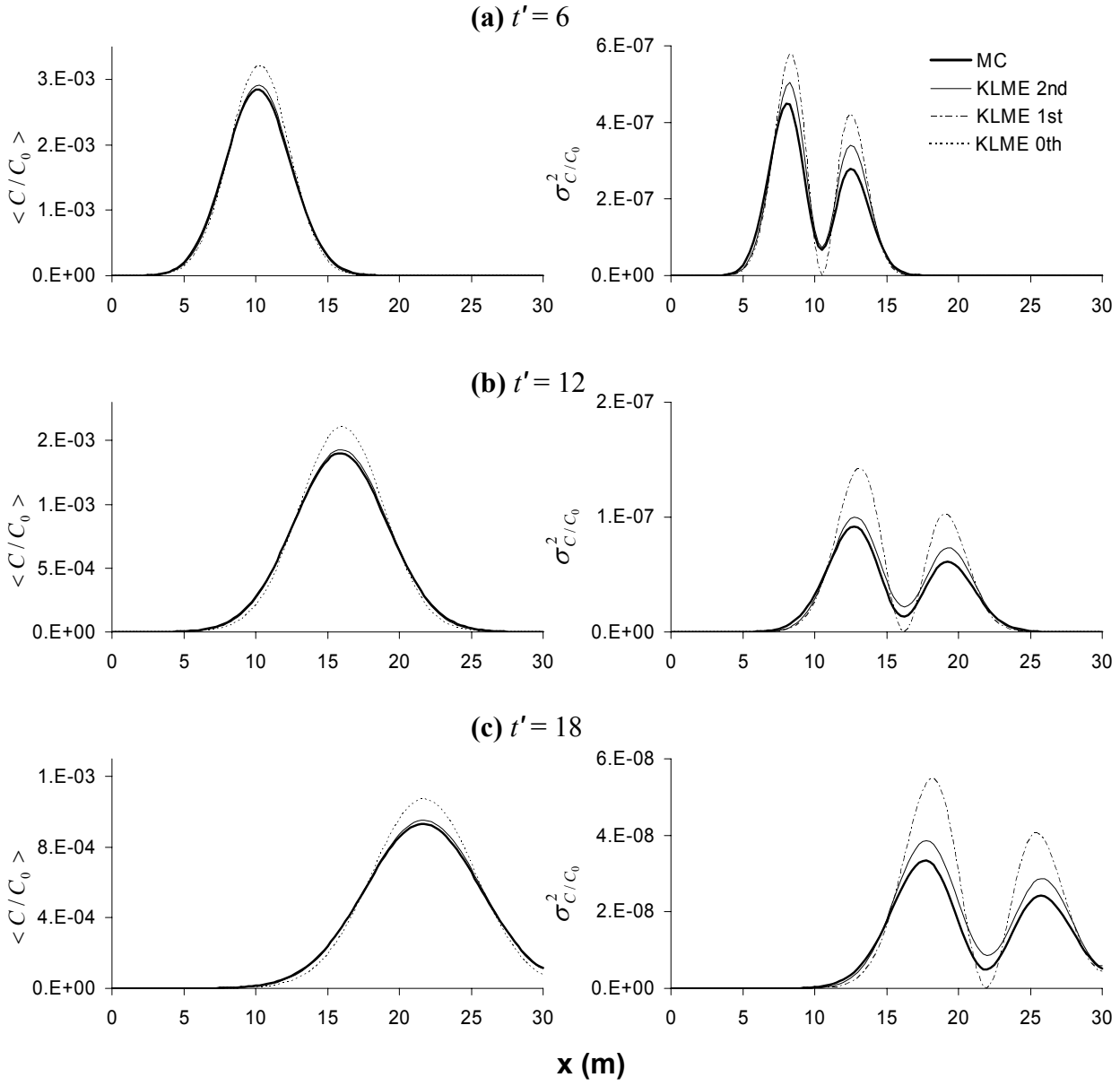


Figure 4. The means and variances of concentrations along the profile A-A' at three different times in Figure 2: **(a)** $t' = 6$, **(b)** $t' = 12$, **(c)** $t' = 18$. The legends for **(b)** and **(c)** are provided in **(a)**.

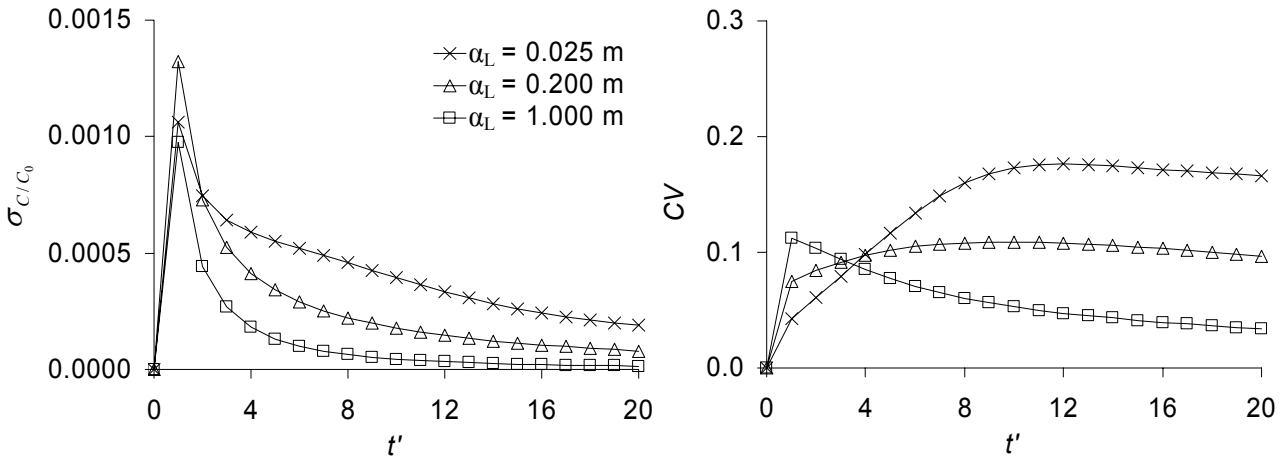


Figure 5. The standard deviation (σ_{C/C_0}) and coefficient of variation (*CV*) at plume centers calculated in the KLME approach at different dispersivities.

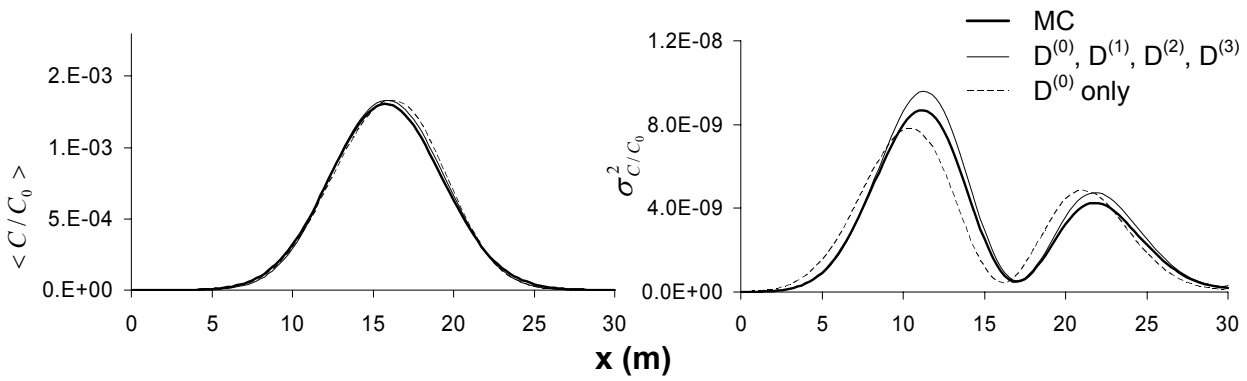
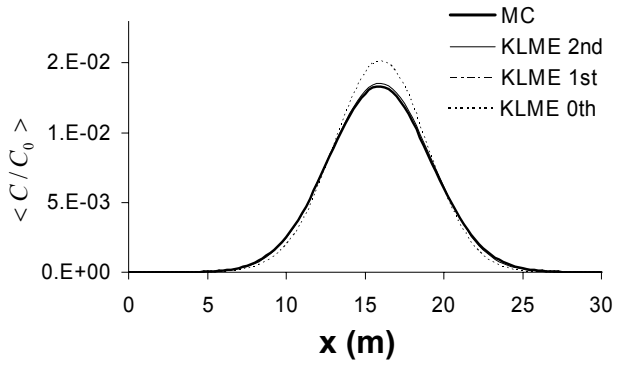
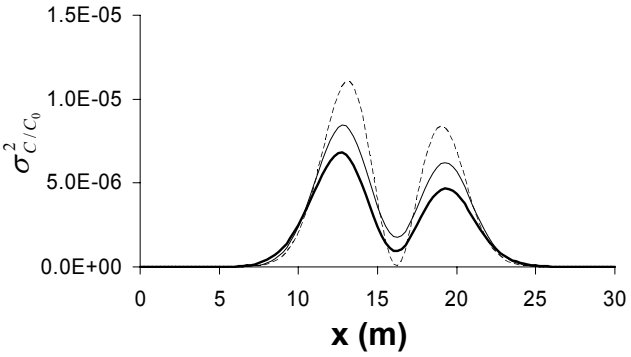


Figure 6. Comparisons of the mean concentrations and concentration variances calculated by the KLME approach between the case where only zeroth-order dispersion term $D^{(0)}$ is used and the case where the higher-order terms $D^{(1)}$, $D^{(2)}$ and $D^{(3)}$ are also included. Results are shown for t' = 12.

(a) $< C / C_0 >$, $t' = 12$



(b) σ^2_{C/C_0}, $t' = 12$



(c) *CV* at Plume Center

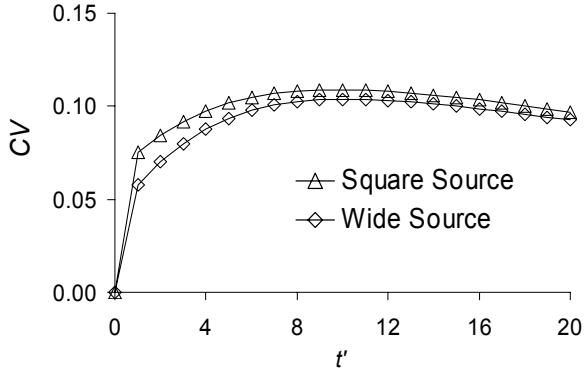


Figure 7. Calculated results in the wide source: (**a**) $< C / C_0 >$ and (**b**) σ^2_{C/C_0} along the profile A-A' at $t' = 12$, and (**c**) *CV* at plume center through different times. The legend for (**b**) is provided in (**a**). The *CV* calculated for the square source is also shown for comparison.

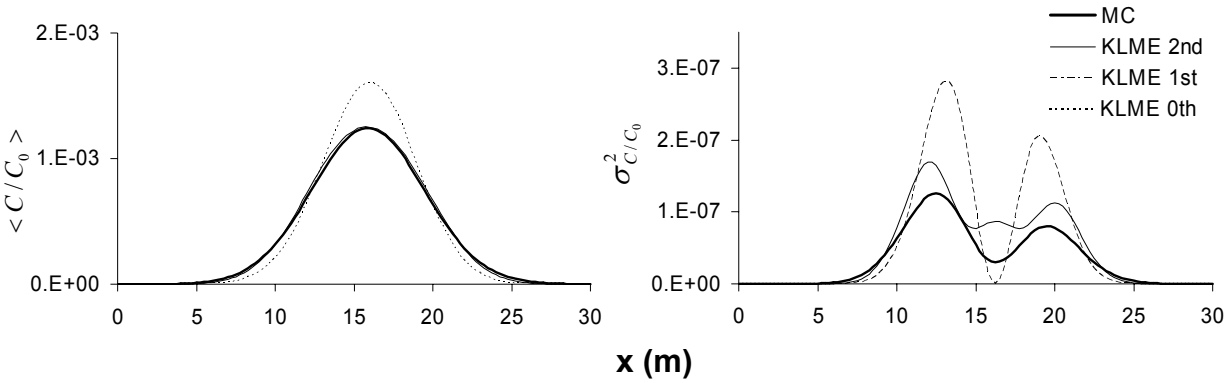


Figure 8. The means and variances of concentrations along the profile A-A' at $t' = 12$ when σ_Y^2 is increased to 0.5.

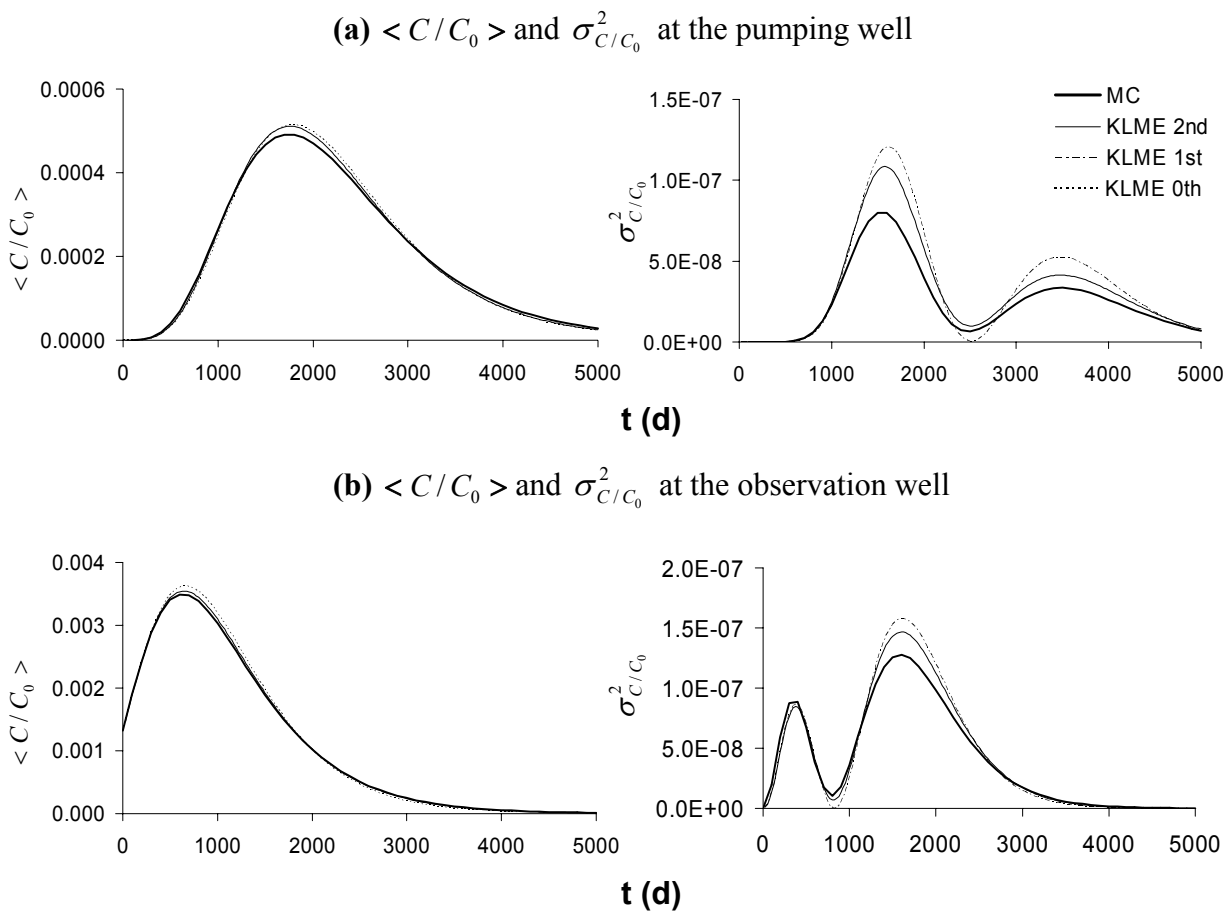


Figure 9. The breakthrough curves for the mean concentrations and concentration variances at (**a**) the pumping well and (**b**) the observation well. The legend for (**b**) is shown in (**a**).